15th International Roundtable on Business Survey Frames
Washington, D.C. – October 22 – 26, 2001

Session No *4*
Paper No 2
Jarmo Ranki*, *Tuula Viitaharju*, *Statistics Finland, Finland

Quality Project of the Business Register

Abstract

Statistics Finland's Business Register belongs to the category of registers that are used both as sampling frames and for producing statistical outputs. As the Finnish statistical production system is very much based on administrative files and registers, the Business Register, too, is widely used as a productional element in other statistical systems - not only for business statistics but also for social statistics. In addition, Business Register data is used for information services subject to charge. Hence, we claim that our Business Register functions as a multipurpose base register.

Lately, the use of the Business Register has intensified to such an extent that the accuracy and correctness of its location data have become more and more important, which is why Statistics Finland has recently made great efforts to improve their quality in this respect.

The paper will deal with the main tasks of the quality project. First, special quality coding has been developed to show and measure the standard of the quality of local unit address data. With the help of this coding we can tell our customers about the quality of our material. Second, a special quality survey has been developed. This survey is monitored according to the quality coding and the questionnaire is kept as simple as possible in order to minimise the response burden of small businesses. Third, new sources are being utilised in the maintenance of address data, namely the data of Finland Post. Fourth, local unit data in the Business Register are being equipped with map co-ordinates. This is done using two methods: initially by linking the data in the Register of Buildings and Dwellings with the local unit data in the Business Register and then by applying geo-coding to equip the rest of the local units with map co-ordinates.

1. Coding of location address quality and present quality of location data in the Business Register

The quality of the location address of a local unit is monitored in the Business Register with a ***coding of location address quality***. The coding examines, e.g. the following:
- When the legal unit was last included in a survey (*the criterion being survey year –2 years)* and whether it responded to it.
- Last changing date the address (other sources)
– Date of verification of the address
– The size of the legal unit *(e.g. turnover or total wagebill)*.

The location addresses of establishments are graded in the quality coding as follows:
– Location address quality is ***good***
– Location address quality is ***probably good***
– Location address quality is ***somewhere between good and poor***
– Location address quality is ***probably poor***
– Location address quality is ***poor***
– Street address is ***missing***
– Location address is of ***defective format*** e.g. street number element is missing or location address is a post office box number.

Even if the coding indicates the quality of the location address of a local unit to be poor, it does not mean that the location address is erroneous as such. The coding of location address quality reveals that the examined location address has not been verified within the last two years (*survey year –2 years)*. The coding thus primarily discloses which location addresses should next be verified. All local units not receiving quality ratings good or probably good for their location addresses are such **potential verification targets**.

Table 1. **Coding of local unit address quality, September 2001 situation.**

Location address coding	**Single establishment legal units**	**Share, %**	**Multiple establishment legal units**	**Share, %**
Good	67,000	20	33,000	85
Potential verification targets	274,000	80	6,000	15
Total	341,000	100	39,000	100

The total number of single establishment legal units is 341,000 of which 274,000 legal units, or 80 per cent, are potential verification targets. The vast majority of these potential verification targets are legal units where the number of salaried employees is under 2 persons. Such legal units number 247,000, in other words 90 per cent of the potential verification targets employ under 2 persons.

Multiple establishment legal units have a total of 39,000 establishments, of which 6,000 local units, or approximately 15 per cent, are potential verification targets.

2. Questionnaire inquiries

In the Finnish Business Register system an annual questionnaire is sent to a legal unit if it has multiple establishments, if it is included in the sample of the survey of single establishment legal units, or if it is a new legal unit.

The annual inquiry of multiple establishment legal units covers close on all multiple establishment enterprises in the country. The inquiry asks for information concerning the activities of both the legal unit and its establishments.

All large legal units and a representative sample of the small legal units where structural changes are suspected to have taken place are selected to the survey of single establishment legal units (Y-*inquiry*). The purpose of this inquiry is to determine whether the units have one or more establishments and discover transfers of establishments, i.e. changes in ownership.

Towards the end of the year 2000, the quality project introduced a questionnaire inquiry directed especially to small, single establishment legal units, or so-called ***Quality Survey*** *(Q-inquiry)*, for verifying changes in location addresses and activity. The selection criteria to this survey take into consideration possible errors in the data on location addresses or type of activity.

Data from new legal units are collected with an inquiry sent to them twice a year. Information about new legal units is obtained from the Tax Authorities. If a new legal unit has activity, it is included in the inquiry.

Table 2. **Questionnaire inquiries of the Business Register**

Questionnaire inquiries	**Target population**	**Units per annum**
Inquiry of multi-establishment legal units	Multi-establishment legal units	- Approx. 8,000 legal units - Approx. 35,000 establishments
Inquiry of single establishment legal units: Y inquiry	All large legal units and sample of the small units where structural change is suspected	- Varying sample size - Approx. 5,000–10,000 legal units
Inquiry of single establishment legal units: Q inquiry	Sample of small legal units	- Varying sample size - Approx. 10,000–20,000 legal units
Inquiry of new legal units	New legal units	- Twice a year - 5,000–8 000 legal units per inquiry round

Changes in the quality codings of the location addresses of establishments were allowed for in the quality inquiries implemented in the year 2000. In other words, when the location address received one of the following quality codings:
– Probably poor,
– Poor,
– Street address missing or
– Defective format,
the pertinent legal unit was included in the sample frame of the quality inquiry.

Two quality inquiries of 10,000 legal units have been conducted in the year 2001. Quality inquiries improve the quality of the data on the structure and location addresses of small, single establishment legal units, so that the distribution of the units' employees by municipality can be reliably ascertained.

3. Data of Finland Post

Because it is not possible to increase the number of the Business Register's own questionnaire inquires any further, other possible sources for location address data have been sought.

Towards the end of 2001, the Business Register will start utilising business address data obtained from Finland Post. These relate to actual visiting addresses and, therefore, agree with the Business Register's concept of establishment location address. In the Business Register, these addresses will primarily be utilised to establish location address data for small, single establishment legal units.

<u>The system for processing Finland Post data</u>

The location addresses of the data of Finland Post are compared with the Business Register's location data on corresponding single establishment legal units:

1) When the location address of Finland Post ***is the same*** as the current location address in the Business Register, the verification date of the location address of the establishment is updated. In the next quality coding updating run, the location address will receive quality coding *good.*

2) The current location address of Finland Post ***is not the same*** as the current location address in the Business Register, but the start date of the current location address of Finland Post is later that that of the current location address in the Business Register. If the criteria for updating and verification are

met, the current location address of Finland Post is updated in the Business Register database as the location address of the establishment. This group includes legal units whose location address is e.g. a post office box number or whose street address is completely missing in the Business Register database. Furthermore, this group comprises single establishment legal units whose location addresses have not been verified for a long time, or never.

The quality of the location data of small, single establishment legal units in the Business Register can be improved notably with the data of Finland Post. The up-to-dateness of the location address data of a considerable number of small, single establishment legal units is verified at the time of the first supply of data. Location address changes, obtained from Finland Post, are subsequently regularly updated in the Business Register.

4. Determining the location municipality of an establishment and the map co-ordinates of the location address

4.1 Co-ordinates

In the Business Register system, the location municipality of an establishment is determined using the address data system of Finland Post. This address data system (*address database*) contains data on the location municipalities and postal code areas of roads and streets. The address database is exploited in the Business Register to verify the existence and consistency of data municipalities, postal code areas, street names and building numbers when establishment location data are being added to, or changed, in the Business Register's production database. If a postal code area spans over a number of municipalities, the system may fail to determine the location municipality correctly.

For this reason a system has been introduced in the Business Register in 2001 whereby the location municipality of the location address of an establishment is also verified against the location data (address, co-ordinate) of the ***Population Register Centre's Register of Buildings and Dwellings (RBD)***. If the verification of the location of an establishment fails using these data, efforts are made to determine the co-ordinate/location municipality of its location address by ***geo-coding*** (locating on a map).

If the location municipality in the address data system of Finland Post and that in the RBD (or the geo-coded location municipality) are the same, we can be quite sure that the location municipality of the location address being verified is correct, and a map co-ordinate for the location address of an establishment is then entered into the Business Register system.

If the location municipality in the address data system of Finland Post and that in the RBD (or the geo-coded location municipality) differ, further verifications must be made before absolutely correct location municipality can be ascertained and a map co-ordinate for the location address of an establishment can be entered into the production database of the Business Register. Once the right location address has been ascertained, the location municipality of the establishment is changed as necessary.

The location address of an establishment has simultaneously only one map-co-ordinate in the Business Register: either the **building co-ordinate** of the RBD or **an address co-ordinate** obtained by geo-coding. The building co-ordinate of the RBD is the primary co-ordinate in the Business Register.

Co-ordinates are entered into the production database of the Business Register in uniform co-ordinate format. A uniform co-ordinate is presented as follows:
y co-ordinate + x co-ordinate, where y = east co-ordinate and x = north co-ordinate.

The x and y co-ordinates are always presented with seven digits (accuracy of one meter). Where the accuracy of a co-ordinate falls below this, noughts are added to its end to extend it to seven digits.

4.2 Population Register Centre's Register of Buildings and Dwellings (RBD)

The Register of Buildings and Dwellings contains data on the locations of buildings: location addresses *(street addresses)* and their co-ordinates. It also contains information on the postal code areas and location municipalities of the location addresses. The co-ordinates in the Register of Buildings and Dwellings are building co-ordinates. The building co-ordinate describes the map co-ordinate of the centroid of a building.

The basic data for the co-ordinates were collected from municipalities in 1970. The data were supplied with the accuracy of 10 metres, although this was extended to 500 metres in a few municipalities in eastern and northern Finland. Continuous maintenance of the data (new buildings and extensions and alterations requiring building permissions) is implemented through municipal building supervision. Since 1980, it has been possible to give the data with the accuracy of one metre.

The co-ordinate data have been verified towards the end of the 1980s and in connection with the 1990 population and dwelling census. Over 360 municipalities have participated in the verifications. The scope and quality of the verification work vary a great deal from one municipality to another. In some municipalities the quality is extremely high and all buildings have co-ordinates that locate to the centroids of the buildings with the accuracy of one metre.

4.3 Geo-coding

Geo-coding *(locating on map)* is an in-built function of the MapInfo and ArcView programs, and facilitates the assigning of map location data to co-ordinate-based data. Statistics Finland's geographic information section locates on the map the location addresses of such establishment for which no co-ordinates are found in the RBD.

The address co-ordinate obtained by geo-coding describes the co-ordinate of the street/road number. It is measured in multiples of ten of metres from the main junction. Distances of street/road numbers from main junctions are measured in fixed units. The further away a building is located from the edge of the road/street, the more its address co-ordinate deviates from its building co-ordinate.

4.4 Different processing stages

1. One-off updating of co-ordinates into the database in February 2001
In the first RBD round, co-ordinates were updated for 300,000 establishments.

2. Obtaining new co-ordinates for changed location addresses
- New RBD co-ordinates are obtained every second week for changed (new) location addresses.
- If no RBD co-ordinates are available, co-ordinates are obtained by geo-coding

3. New RBD version
The co-ordinates of the current location addresses of all establishments in the production database are verified against the new RBD version, so that all co-ordinate verifications and corrections can be entered into the production database.

Differences/changes between the RBD versions of different years:
- New buildings
- Removed buildings
- Accuracy of co-ordinates increases, e.g. from 100 metres to 10 metres
- Erroneous co-ordinates are corrected.

4. Old co-ordinates are saved as historical data
When the location address of an establishment changes, the old address and its co-ordinate are saved as historical data.

5. Additional clarification of the location municipality of an establishment
No RBD co-ordinates are entered into the production database without further verification for such establishments for which the location municipality in the RBD/location municipality obtained by geo-coding differs from the location municipality in the production database. In such cases, separate investigations are conducted to ascertain the correct location municipality of the establishment concerned. Once the correct location municipality has been established, the location of municipality of the establishment is changed in the production database as necessary.

4.5 Number of co-ordinates in the production database of the Business Register

Co-ordinates have been sought for all the establishments that have a record for the year 2000 in the production database of the Business Register and for establishments that have ceased to operate during the year 2000. The size of this base population is 460,000 establishment. No co-ordinates are sought for establishments having ceased operations in the year 1999 or prior to it.

Thus far, the building co-ordinates of 350,000 establishments have been updated in the production database. These 350,000 establishments have exactly matching location municipalities in both the address database of Finland Post (= current location municipality of the production database) and in the RBD.

A total of 8,000 establishments have been found in RBD co-ordinate runs for which the current location municipality in the production database and that in the RBD differ from each other. Further verifications will be done in respect of these establishments to ascertain the correct location municipality.

No co-ordinates have so far been found for good 100,000 establishments:
– The location address record (*street address data*) is empty in respect for 28,000 establishments in the production database of the Business Register.
– No co-ordinates can be found in the location address record of the RBD in respect of 73,000 establishments and the location address of 33,000 of these establishments is invalid in the production database of the Business Register *(e.g. street number element is missing or the location address is a post office box number).*

The first geo-coded data are also being processed at the moment and good 6,000 address co-ordinates obtained by geo-coding will soon be updated into the production database of the Business Register.